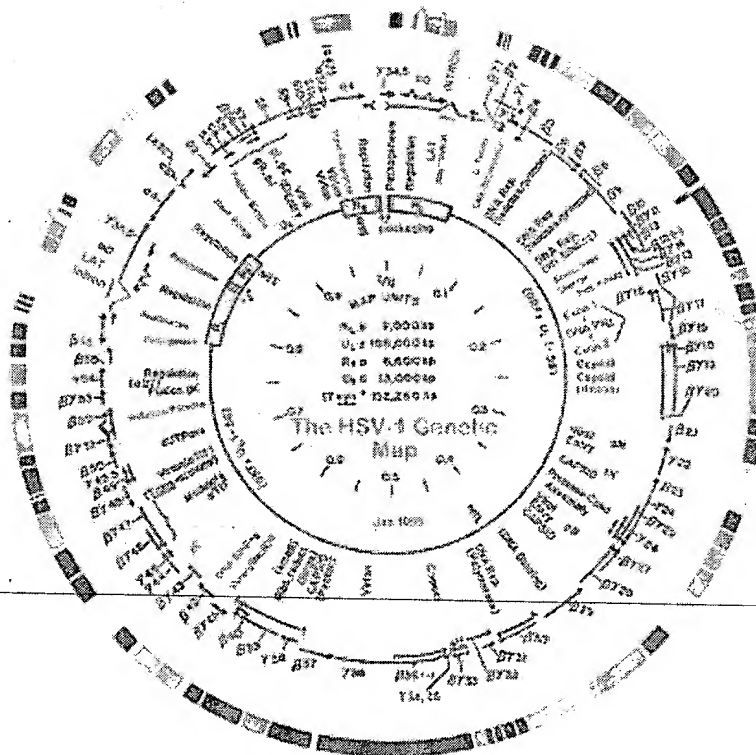


HSV DNA - ca. 47 μ m length

The HSV-1 genome is a linear, double stranded DNA duplex 152,000 base pairs in length, and with a base composition of 67% G+C. The genome circularizes upon infection. Because the genome circularizes, the transcription and genetic map is conveniently shown as a circle.



Since the virus encodes nearly 100 transcripts and more than 70 open translational reading frames (ORFs), the map is complex. Still, lack of splicing of HSV transcripts means that most

ORFs are expressed by a single transcript.

The HSV genome can be divided into six important regions.

- 1 The ends of the linear molecules, which are important in both circularization of the viral DNA upon infection, and in packaging the DNA in the virion. The DNA replicates by a type of rolling circle mechanism.
 - 2 The 9,000 bp long repeat (R_L), which encode both an important immediate early regulatory protein ($\alpha 0$) and the promoter of and most of the "gene" for the latency associated transcript (LAT).
 - 3 The long unique region (U_L), which is 108,000 bp long, encodes at least 56 distinct proteins (actually more because some ORFs are spliced and expressed in redundant ways). It contains genes for the DNA replication enzymes and the capsid proteins, as well as many other proteins.
 - 4 The 6,600 bp short repeats (R_S) encode the very important $\alpha 4$ immediate early protein. This is a very powerful transcriptional activator. It acts along with $\alpha 0$ and $\alpha 27$ (in the U_L) to stimulate the infected cell for all viral gene expression that leads to viral DNA replication.
 - 5 The origins of replication. The ori_L is in the middle of the U_L region. The ori_S is in the R_S and, thus, is present in two copies. All sets of ori's operate during infection to give a very complicated replication complex.
 - 6 The 13,000 bp unique short region (U_S) encodes 12 ORFs, a number of which are glycoproteins important in viral host range and response to host defense.
-